

DAFTAR PUSTAKA

- [1] B. Liu, *Web Data Mining : Exploring Hyperlink, contents, and Usage Data*. Spriger, 2007.
- [2] A. Huang, "Similarity measures for text document clustering," in *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008)*, Christchurch, New Zealand, 2008, pp. 49–56.
- [3] X. Wang, J. Cao, Y. Liu, S. Gao, and X. Deng, "Text Clustering Based on the Improved TFIDF by the Iterative Algorithm," in *IEEE Symposium on Electrical & Electronics Engineering (EESYM)*, 2012.
- [4] A. Guo and T. yang, "Research and Improvement of feature words weight based on TFIDF Algorithm," in *IEEE*, 2016.
- [5] W. Yong-qing, L. Pei-yu, and Z. Zhen-fang, "A Feature Selection Method based on Improved TFIDF," in *IEEE*, 2008.
- [6] M. Tutkan, M. C. Ganiz, and S. Akyokuş, "Helmholtz principle based supervised and unsupervised feature selection methods for text mining," *Inf. Process. Manag.*, vol. 52, no. 5, pp. 885 – 910, 2016.
- [7] W. H. Gomaa and A. A. Fahmy, "A Survey of Text Similarity Approaches," in *International Journal of Computer Applications (0975 – 8887)*, 2013.
- [8] K. P. N. V. S. sree and J. V. R. Murthy, "Clustering Based on Cosine Similarity Measure," in *International Journal Of Engineering Science & Advanced Technology*, 2012.
- [9] R. K. Roul, O. R. Devanand, and S. K. Sahay, "Web Document Clustering and Ranking using Tf-Idf based Apriori Approach," 2014.
- [10] A. Fuddoly, J. Jaafar, and N. Zamin, "Keywords Similarity Based Topic Identification for Indonesian News Documents," in *European Modelling Symposium*, 2013.
- [11] Alexander Strehl, Joydeep Ghosh, and Raymond Mooney, "Impact of Similarity Measures on Web-page Clustering," in *AAAI Technical Report WS-00-01*, 2000.
- [12] J. A. Kumar and S. Abirami, "An Experimental Study Of Feature Extraction Techniques In Opinion Mining," in *International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI)*, 2015, vol. 4.
- [13] C. Su, Q. Chen, X. Wang, and X. Meng, "Text Clustering Approach Based on Maximal Frequent Term Sets," in *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*, 2009.
- [14] Aini Rahmania, J. Jaafar, and N. Zamin, "Likelihood calculation classification for Indonesian Language News Documents," in *IEEE*, 2013.
- [15] Petr Kroha and Ricardo Baeza-Yates, "A Case Study: News Classification Based on Term Frequency," in *Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA'05)*, 2005.
- [16] I. H. Witten, E. Frank, and M. a Hall, *Data Mining : Practical Machine Learning Tools and Techniques (Google eBook)*. 2011.
- [17] Manning, C. D., Raghavan P. and Schutze, H., "An Introduction to

- Information Retrieval,” Cambridge, England, 2009.
- [18] Feldman, R. and Sanger, J., *The Text Mining HandBook: Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge University Press, 2007.
 - [19] Adriani, M., Asian, J., Nazief, B., Tahaghoghi, S. M. M., and Williams, H. E., “Stemming Indonesian,” in *ACM Transactions on Asian Language Information Processing*, 2007, vol. 6, pp. 1–33.
 - [20] Bobby A. A. Nazief and Mirna Adriani, “Confix Stripping: Approach to Stemming Algorithm for Bahasa Indonesia,” in *Internal publication*, Faculty of Computer Science, University of Indonesia. Depok.
 - [21] JONES K S, “A statistical interpretation of term specificity and its application in retrieval[J],” in *Journal of Documentation*, 1972, vol. 28(1), pp. 11– 21.
 - [22] SALTON G and CLEMENT T Y, “On the construction effective vocabularies for information retrieval[C],” in *Proceedings of the 1973 Meeting on Programming Languages and Information Retrieval.*, New York, 1973.
 - [23] “Silhouette (clustering),” *Wikipedia*. [Online]. Available: [https://en.wikipedia.org/wiki/Silhouette_\(clustering\)](https://en.wikipedia.org/wiki/Silhouette_(clustering)). [Accessed: 19-Jul-2017].
 - [24] Marco Bonzanini, in *Mastering Social Media Mining with Python*, Packt Publishing Ltd., 2016.
 - [25] L. Kaufman and P. J. Rousseuw, *Finding Groups in Data*. New York: John Wiley & Sons, 1990.